

# SEQUENCE LISTING

<110> ROULEAU, Guy A.  
 JOOBER, Ridha  
 BENKELFAT, Chawki

<120> POLYMORPHIC CAG REPEAT-CONTAINING GENE AND USES THEREOF

<130> 12948-10

<140> 09/508,821

<141> 2000-05-26

<150> PCT/CA98/00884

<151> 1998-09-18

<150> 2,216,057 CA

<151> 1997-09-19

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 328

<212> DNA

<213> Mus musculus

<400> 1

```

agggcagcca ctttccccag cattccccagt ccttccttac ctctccact tatgccccaa 60
cagtgcaggg tggtagggcag ggggcccact cctacaagag ctgcacagca ccatctgccc 120
agcctcatga taggccgatg agtgccaatg cgaacctggc tccagggcaa cgggtccaga 180
atcttcacgc ttaccagcct ggccgccttg gctacgagca gcagcagcaa gcacttcaag 240
gccgtcacca caccaggaa aactccact accagaacct cgccaagtac caacactatg 300
gacagcaagg ccagggtac tgtccacc 328
  
```

<210> 2

<211> 330

<212> DNA

<213> Homo sapiens

<400> 2

```

tccttcccc cctcctccac ctactcctcc tctgtcccag ggtggtgggc agggggccna 60
ctcctataag agttgcacag caccgactgc ccagcccat gacaggccgc tgactgccag 120
ctccagcctg gccccggggc agcgggtcca gaatcttcat gcctaccagt cgggcgcgct 180
cagctatgac cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca 240
gagccggcac catgcccagg aaaccctcca ttaccaaaac ctgcccaagt atcagcacta 300
  
```

RECEIVED

OCT 17 2000

TECH CENTER 1606/2900

RECEIVED

OCT 26 2000

TECH CENTER 1606/2900



a

cgggcagcaa ggccagggct actgccagcc

330

<210> 3

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
oligonucleotide

<400> 3

ggggcagcgg gtccagaatc ttc

23

a<sup>1</sup>  
<210> 4

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
oligonucleotide

<400> 4

agcactacgg gcagcaaggc cag

23

<210> 5

<211> 6022

<212> DNA

<213> Homo sapiens

<400> 5

ggatccagca ggcccaagg gatgagggag cggaaattgc tctgctaaat gcttttgagc 60  
tgtcaggaag ggctgggagt gatgggtggg ggacattggg gaggagctgg caatgggcgg 120  
ggggggggcg ggtagctccc cagtgcctg gcgctgggca gccggttttg cctcccgcac 180  
cagtggccgt ccttggaag actcagctgc agcgatgtg ggagcgggaa ttacagagca 240  
cacctccctg acacagaagt tgtcaatatg cgcacagctg gtggggaggc tcaggcgaag 300  
gggggactat taagagctgc gcgggggagc aggcagggtg gggaggtggg tgggaggggtg 360  
ctttctgagg caaaaggaag tggcccgctc gaatcgctca tcctctgccc cctccctgcc 420  
catectcccc tccctccttc cctccctccc tcccttcctt tttcttttca cagataacca 480  
gcccagagtca tgcagtcttt tcgagaaagg tgtggtttcc atggcaaaca acagaactac 540  
cagcagacct cgcaggaaac atcacgccta gagaattaca ggcagccgag tcaggccggg 600  
ctaagctgag accggcagcg gctgctcgcc aaggactatt ataaccgca gccttaccgg 660  
agctatgagg gtggcgctgg cagccctct ggcaactgag ccgcggtggc cgccgacaag 720  
taccaccgag gcagcaaggc cctgcccaca cagcaaggcc tgcaaggggag gccggctttc 780

cctggctacg gcgtccagga cagcagcccc taccaggcc gctatgctgg tgaggagagc 840  
cttcaggctt ggggggcccc acagccacca cccccacagc cgcagccact acctgcaggg 900  
gtggccaagt atgatgagaa cttgatgaaa aagacagcag tgccccccag caggcagtat 960  
gcagagcagg gcgcccaggt gcccttttcgg actcaactccc tgcacgtcca gcagccaccg 1020  
ccgccccagc agccccctggc ataccccaag ctccaaaggc agaagctgca gaacgacatt 1080  
gcctccccctc tgcccttccc ccagggtacc cacttttcctc agcattccca gtccttcccc 1140  
acctcctcca cctactcctc ctctgtccag ggtggtgggc agggggccca ctccataaag 1200  
agttgcacag caccgactgc ccagcccat gacaggccgc tgactgccag ctccagcctg 1260  
gccccggggc agcgggtcca gaattctcat gcctaccagt cggggccgcct cagctatgac 1320  
cagcagcagc agcagcagca gcagcagcag cagcagcagc aagcccttca gagccggcac 1380  
catgcccagg aaaccctcca ttaccaaaac ctgcgaagt atcagcacta cgggcagcaa 1440  
ggccagggtc actgccagcc ggacgcagcc gtccggaccc cagagcagta ctaccagacc 1500  
ttcagcccca gctccagcca ctaccccgcc cgctccgtgg gccgtcacc ttcctacagt 1560  
tccacaccgt cgccgtgat gccaaacctg gagaactttc cctacagcca gcagccgtc 1620  
agcaccgggg ccttccccgc agggatcact gaccacagcc acttcatgcc cctgtcaat 1680  
ccctcccaa cggatgccac cagctctgtg gacaccagc ctggcaactg caagccctt 1740  
cagaaggaca agctccctga gaacctgtg tcggatctca gcctgcagag cctcacggcg 1800  
ctgaccttac aggtggagaa catctccaac accgtccagc agctgtgtgt ctccaaggct 1860  
gctgtgccgc agaagaaagg tgtcaagaac ctggtgtcca ggacccaga gcagcataaa 1920  
agccagcact gcagccccga agggagcggc tactcagccg agcccgagc cacaccgtg 1980  
tcagagccgc cgagcagcac gccacagtc acgcatgcgg agccgcagga ggccgactac 2040  
ctgagcggct ccgaggaccc actggagcgc agcttctct actgcaacca ggcccggtgg 2100  
agccctgcca ggggtcaacag caactcgaag gccaaagccg agtccgtgtc cacctgttct 2160  
gtgacctctc ctgacgacat gtccaccaa tctgacgact ccttcagag cctacacggc 2220  
agtctgccgc tcgacagctt ctccaagttc gtggcgggtg agcgggactg tccgcggctg 2280  
ctgctcagcg ccctggcaca ggaggacctg gcctccgaga tcctggggct gcaggaagcc 2340  
atcggtgaga aggcgcagaa agcttgggtc gaagcaccca gcctggtcaa ggacagcagc 2400  
aagccacctt tctcgttga gaaccacagc gcctgcctgg actctgtggc caagagtgcg 2460  
tggccccggc ctggggagcc ggaggccctg cccgactcct tgcagctgga caagggcggc 2520  
aatgccaaag acttcagccc agggctgttt gaagaccctt ccgtggcctt cgctacgct 2580  
gacccccaaa agacaactgg tcctctctcc tttggtacca agccccctt tggggttcct 2640  
gctccagacc ccactacagc agcttttgac tgtttccgg acacaaccgc tgccagctca 2700  
gcggacagcg ccaaccctt tgctggcca gaggaacc tgggggatgc ttgtccagg 2760  
tggggattgc accctggcga gcttaccaag ggctggagc aggggtggaa ggctcagat 2820  
ggcatcagca aaggggacac ccattgaggt tcggcctgcc tgggcttcca ggaggaggac 2880  
ccccctggg agaaggtggc ctctgtgccc ggggacttca agcaggagga ggtgggtggg 2940  
gtgaaggagg aggcaggtgg gctgctgcag tgccccgagg tggccaaggc tgaccggtgg 3000  
ctggaggaca gccggcactg ctgttccacc gccgacttcg gggacctccc actgtgcca 3060  
cccaccagca ggaaggagga cctggaagct gaggaggagt actcctcct atgtgagctc 3120  
ctgggcagcc ccgagcagag gcctggcatg caggaccgc tgtcacccaa ggccccactc 3180  
atctgcacca aggaggaggt ggaggaggtg ctggactcca aggcgggtg gggctctccg 3240  
tgccacctct caggggagtc cgtcatcctg ctgggcccta cagtgggcac cgagtcaaag 3300  
gtccagagct ggtttgagtc ctctctgtca cacatgaagc caggtgaaga ggggctgat 3360  
ggggagcgag ctccagggga ttccaccacc tcggacgct ctctggccca gaagcccaac 3420  
aagcctgtg tgcccgaggc gcccatcgca aagaaagagc ctgtgccacg gggcaaaagc 3480  
ttacggagcc gtcgggtgca ccgggggtg cccgaggccg aggaactccc atgcagggca 3540  
ccagtgtgct ccaaagacct ctgtctcct gaatcctgca cagggccccc ccagggacag 3600  
atggaagggg ctggagcccc aggcggggg gcctcggaag ggctccccag gatgtgtact 3660

a

cgttctctca cgccctgag tgagcccg acgcccggac ccccaggcct gaccaccacc 3720  
 cctgcacccc cagacaaact ggggggcaag cagcgagccg ccttcaagtc gggcaagcgg 3780  
 gtggggaagc cctcacccaa ggctgcctcc agcccagca acccgccgc cctgcctgtg 3840  
 gcctccgaca gcagcccgat gggctccaag accaaggaga cagactcacc cagcacgcct 3900  
 ggcaaggacc agcgctccat gatccttcgg tcacgcacca aaaccagga gatcttcac 3960  
 tccaagcggc ggaggccctc tgaggcgcg ctccccaact gccgtgccac caagaagctc 4020  
 ctcgacaaca gccacttgcc cgccacattc aaggtctcca gcagcccca gaaggagggc 4080  
 aggggtgagcc agcgggcaag ggtcccaaaa cctggtgcag gcagcaagct ctctgaccgg 4140  
 cccctccatg cgctcaaaag gaagtcggcc ttcattggcg cgggtccccc caagaagcgg 4200  
 aacctggtct tgcggcacgg cagcagcagc agcagcaacg ccagtgcatt gggggagatg 4260  
 ggaaggagga gaggcctgag ggttcccca cctcttcaa gaggatgtct tctccaaga 4320  
 aagccaagcc caccaagggc aatggcgagc ctgccacaaa gctccacccc ccggagaccc 4380  
 ccattcctgc ctcaagctcg cctctcggca gccttcagg gggccatgaa gaccaaggtg 4440  
 ctgccacccc ggaaggggcg gggcctgaag ctggaagcca tcgtgcagaa gatcacctcg 4500  
 ccagcctca agaagttcgc atgtaaagcg ccaggggcct ctctggtaa tctctgagc 4560  
 ccatcccttt ccgacaaaga ccgtgggctc aagggtgctg ggggcagccc agtgggggtg 4620  
 gaagaaggcc tggtaaagt gggcaccggg cagaagctcc caactcttg ggctgatccg 4680  
 ttatgcagaa atccaaccaa cagatcctta aaaggcaaac tcatgaacag taagaaactg 4740  
 tcttctactg actgtttcaa aaccgaggcc ttcacatccc cggaggccct gcagcctggg 4800  
 gggactgcc tggcgccaa gaagaggagc cggaaaggcc gggcaggggc ccatggactc 4860  
 tccaaaggcc cgctggagaa gcggccctat cttggcccgg ctctgtcct gactccccga 4920  
 gacagggcca gtggcacaca aggggccagt gaggacaact ctggtggagg aggcaagaag 4980  
 ccaaagatgg aggagctgg ccctgcctcc cagcccccg agggcaggcc ctgccagccc 5040  
 cagacaaggg cacagaaaca gccaggccac accaactaca gcagctattc caagcggaag 5100  
 cgcctcactc ggggccgggc caagaacacc acctcttcac cctgtaaggg gcgtgccaag 5160  
 cgacgacgac aacagcaggt gctgccctg gatcccgag agcctgaaat ccgcctcaag 5220  
 tacatttct cttgcaagcg gctgaggtca gacagccgga ccccgccct ctcacccttc 5280  
 gtgcgggtgg agaagcgaga cgcgttcacc accatatgca ctgttgtaa ctcccctgga 5340  
 gatgcgcca agccccacag gaagccttc tctctgct cctcttctc atcctcgtcc 5400  
 tcgttctct tggatgcagc cggggcctcc ctggccacac tccctggagg ctccatcctg 5460  
 cagccgcggc cctccttgcc cctctcctcc acgatgcact tggggcctgt ggtttccaag 5520  
 gccctgagta cctcttgct tgtttgctgc ctctgcaaaa acccgccaa cttcaaggac 5580  
 cttggggacc tctgtgggccc ctactacct gaacactgcc tccccaaaaa gaagccaaaa 5640  
 ctcaaggaga aggtgcggcc agaaggcacc tgtgaggagg cctcgctgcc gcttgagaga 5700  
 aactcaaaag gtcccagtg tgcagctgcc gccactgcc ggaagcccc caggtgacgg 5760  
 cccagctgac ccggccaagc agggcccaact gcgcaccagt gcccggggcc tgtcccggag 5820  
 gctgcagagc tgctactgct gtgatggccg ggaggatggg ggcgaggagg cagccccagc 5880  
 cgacaagggt cgcaaacatg agtgagcaa ggaggctccg gcagagccc gcggggaggg 5940  
 ccaggagcac tgggtgcatg aggcctgtgc cgtgtggacc ggcggcgtct acctggtggc 6000  
 cgggaagctc ttgggctgc ag 6022

<210> 6

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

a' <223> Description of Artificial Sequence:  
oligonucleotide

<400> 6  
tggccttgct gcccgtagtg ct

=> d his

(FILE 'HOME' ENTERED AT 14:57:30 ON 20 DEC 2001)

FILE 'MEDLINE, BIOSIS, CAPLUS' ENTERED AT 14:57:36 ON 20 DEC 2001  
L1 2044 S ORFX OR GSGT1 OR KIAA1820 OR HGT1 OR GT1 OR RAI1  
L2 170 S L1 AND (SCHIZOPHRENIA OR MANIC OR DEPRESS? OR DISEASE)  
L3 12 S L1 AND CAG  
L4 6 DUP REM L3 (6 DUPLICATES REMOVED)  
L5 170 S L1 AND L2  
L6 128 DUP REM L5 (42 DUPLICATES REMOVED)  
L7 5 S L6 AND SCHIZOPHRENIA  
L8 5 S L7 NOT 4